In AVs We Trust: Conceptions to Overcome Trust Issues in Automated Vehicles

Kai Holländer[1]

***Abstract*— To take advantage of the full potential of highly automated vehicles (AVs), users need to trust the system enough to be willing to engage with the novel technology. In this context, users of automated vehicles face the challenge of understanding the system capabilities. While interacting with an AV users need to calibrate between overtrust (trusting the vehicle beyond its capabilities and underestimation of the consequences if the system fails) and undertrust (not relying on the vehicle even though it is capable of handling the situation perfectly well). For this work I look at crucial aspects which should be considered for the calibration of trust, such as: proper training, appropriate user interfaces and how to possibly measure trust. I believe it is very important that users of AVs understand the systems capabilities and calibrate their trust accordingly. The findings and ideas of this work are mainly based on previous work which has been revised regarding trust in AVs. The most important takeaway is that users of AVs need specific training and a mental model that differs from concepts for manually driven vehicles.**

I. INTRODUCTION

The forthcoming introduction of autonomous vehicles (AVs) might reduce the overall frequency and severity of accidents caused by human failure [1]. At the same time, new challenges and concerns arise, for example the design interfaces for passengers, as well as for external interaction with other cars, pedestrians, or cyclists [12], [13], [25]. Overall, it has been reasoned that AVs demand an all-new way of thinking about corresponding user interfaces design [2], for example with regard to vehicle-user collaboration [3], [4], external communication [5], [6], [7], [25], interior design [8], [9], take-over requests [10] and modalities used [11].

The questions addressed in this work are: (1) *Which aspects influence trust in AVs?* And (2) *How can trust in AVs be calibrated in a meaningful way?*

Discussing and addressing these issues might facilitate the introduction and seamless acceptance of autonomous vehicles for users and potential stakeholders and, in the long run, help to establish the benefits of this new technology in road transport.

In order to present insights into these questions I take a closer look at relevant related works and identify a few (non-exhaustive)several key issues which are open for discussion.

This work presents ideas on how to promote user confidence in AVs, with particular attention to users who have no previous touch points or experience with AVs.

The specific outcome of this investigation is a set of three overarching research opportunities which should be prioritized on the basis of their expected influence on user confidence and possibly be implemented in future pilot studies to find out if further investigations might benefit the calibration of users' trust in automated driving features.

In the long term I hope that the presented thoughts provoke novel prototypes which ease the introduction of AVs for inexperienced users. Furthermore, the presented three research opportunities could be expanded and should be considered as a thought-provoking pillars (and not a complete list).

II. BACKGROUND & RELATED WORK

Trust in automation could become one of the major barriers for a successful market introduction of automated vehicle technology [16], [17]. However, findings on trust in automated driving systems seem to be diverse. Some studies suggest that there are initial trust issues [15] in driving automation technology. One reason for this could be recorded accidents with AVs [18]. On the other hand, there is also user feedback which shows high levels of initial trust [7]. Hence, the question arises: *Which aspects could influence trust in automated driving features?*

Tenhundfeld et al. [14, p1] state that

> "the largest influences on trust in automation is the familiarity with the system."

Currently we cannot assume that many potential users are familiar with fully automated driving systems because the technology is not established in everyday-traffic yet (as of March 2021). Therefore, special considerations should be afforded to inexperienced users. Another core aspect of adequate trust in AV features is the calibration of trust [19], [20], [26]. To this end, three factors must be appropriately Weighed: (1) risks in interacting with a system, (2) consequences of a system failure, and (3) the capabilities of an automated system. Important aspects thereof are adequate and usable user interfaces catering to the communication with vehicle occupant(s) [21], [22], [30] and the external communication with other road users [18], [24]. For the calibration of trust, user interfaces which foster the understanding of the wider context in which the automated vehicle is operated could become an essential basis for a functioning cooperation between drivers and AVs. Furthermore, users of an automated system should not lose their Situational Awareness (SA), otherwise they might commit errors more frequently [23]. According to Endsley [23], [29] SA in the context of dynamic systems is the perception of the elements in an environment bounded by space and time, as well as the understanding of their meaning and the projection of their state in the near future.

[1]Kai Holländer, Faculty for Statistics, Mathematics and Computer Science, LMU Munich, 80337 Munich, Germany kai.hollaender@ifi.lmu.de

Other important aspects which influence trust issues in automated driving could be: personal experience [31], cultural differences [28], overtrust [33] overreliance [33] and age related requirements [27], [31]. For a well-working human computer interaction in automated driving, a suitable mental model of AVs which allows users to create a useful representation of the systems in their consciousness could have a genuine practical influence [34]. For this purpose, Hailong Liu and Toshihiro Hirakoka present a model which could help to prevent overtust in automated driving systems [36].

III. Research Opportunities & Research Recommendations

Based on the mentioned background, I identified three critical research opportunities which are supposed to merit further discussion on their eligibility and eventual further investigation. I did not rank or triage them on purpose, because I conjecture that the priority and importance is dependent on the ongoing development and availability of AVs. I want to strongly emphasize that these are only three factors which I think are crucial for a successful trust calibration in the context of automated vehicles. There are more (e.g. an overall holistic trust through personal experience and emotional connection of users with their AVs). However, for this workshop I want to focus on these three aspects.

- **Provide adequate training for future users**
  This could include mandatory lessons in relation to: misuse, disuse, overtrust, calibration of trust, system capabilities, the mental model of vehicle automation, and user interfaces. This can also encompass country specific regulations.
- **Design appropriate user interfaces:**
  The interfaces of AVs should foster the calibration of trust, situational awareness, and familiarity.
- **Find suitable methods to measure trust**
  Trust could be measured by means of: open interviews and/or questionnaires, such as System Trust Scale (STS) or Subjective Mental Effort Questionnaire (SMEQ) or NASA-TLX [32], [35]. However, given the fact that questionnaire-based metrics of trust are subjective and not real-time, there is a need to identify a more real-time, 'hands-off' way of measuring trust through surrogate measures of physiological responses. Research is needed to determine if specific physiological responses measured by means of e.g. electroencephalography (EEG) or galvanic skin response (GSR), or combination thereof, can provide a workable estimate of user trust, which can subsequently be used to feed the user interfaces or adjust the behavior of the AV accordingly [32].

IV. Outlook & Discussion

I argue that user training, the design of matching user interfaces (which clearly indicate the systems boundaries) and measuring trust are three major concerns which should be addressed in future work in order to create innovative solutions for safe and usable automated vehicle technology.

Furthermore, automated vehicle technology defines a human-machine interface combining psychology and technology, and therefore an adequate mental model should be conveyed to potential users [34]. Specifically, a validated, comparable, and scientifically stable measurement method for reliance and confidence in automated vehicles should be created.

I would like to bring these discussion points into the workshop to spark a conversation regarding the next steps of research into trust in automated driving.

V. Acknowledgements

I especially thank Ingrid Hamm for her support in the central theme of this work. In addition I thank Renate Gasiorek and Debargah Dey for proofreading and improving this work.

References

[1] Litman, Todd. Autonomous vehicle implementation predictions. Victoria, Canada: Victoria Transport Policy Institute, 2017.

[2] Kun, Andrew L., Susanne Boll, and Albrecht Schmidt. "Shifting gears: User interfaces in the age of autonomous driving." IEEE Pervasive Computing 15.1 (2016): 32-38.

[3] Walch, Marcel, et al. Cooperative Overtaking: Overcoming Automated Vehicles' Obstructed Sensor Range via Driver Help. In: Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications. 2019. S. 144-155.

[4] Wiegand, Gesa and and Holländer, Kai and Rupp, Katherina and Hussmann, Heinrich. Strong Together: A Collaborative Driver Interface to Overcome Limitations of Automated Vehicle Systems. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2020. DOI:10.1145/1122445.1122456.

[5] Dey, Debargha, et al. Distance-dependent ehmis for the interaction betIen automated vehicles and pedestrians. In: 12th international conference on automotive user interfaces and interactive vehicular applications. 2020. S. 192-204.

[6] Ranashinge, Champika, et al. Autonomous Vehicle-Pedestrian Interaction Across Cultures: Towards Designing Better External Human Machine Interfaces (eHMIs). In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020. S. 1-8.

[7] Holländer, Kai; Wintersberger, Philipp; Butz, Andreas. Overtrust in external cues of automated vehicles: an experimental investigation. In: Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications. 2019. S. 211-221.

[8] Wiegand, Gesa, et al. Incarar: A design space towards 3d augmented reality applications in vehicles. In: Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications. 2019. S. 1-13.

[9] LI, Jingyi, et al. An Exploration of Users' Thoughts on Rear-Seat Productivity in Virtual Reality. In: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2020. S. 92-95.

[10] Holländer, Kai; Pfleging, Bastian. Preparing drivers for planned control transitions in automated cars. In: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia. 2018. S. 83-92.

[11] Politis, Ioannis; Brewster, S.; Pollick, F. The Effects of Modality, Urgency and Message Content on Responses to Multimodal Driver Displays. AutomotiveUI EA, 2014.

[12] Holländer, Kai. Applying the user-centered design process to external car displays. Mensch und Computer 2018-Workshopband, 2018.

[13] Nguyen, Trung Thanh, et al. Designing for projection-based communication betIen autonomous vehicles and pedestrians. In: Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications. 2019. S. 284-294.

[14] Tenhundfeld, Nathan L., et al. Calibrating trust in automation through familiarity with the autoparking feature of a Tesla Model X. Journal of cognitive engineering and decision making, 2019, 13. Jg., Nr. 4, S. 279-294.

[15] Lee, John D.; Kolodge, Kristin. Understanding attitudes towards self-driving vehicles: Quantitative analysis of qualitative data. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage CA: Los Angeles, CA: SAGE Publications, 2018. S. 1399-1403.
[16] Häuslschmid, Renate, et al. Supporting Trust in autonomous driving. In: Proceedings of the 22nd international conference on intelligent user interfaces. 2017. S. 319-329.
[17] Noah, Brittany E., et al. First workshop on trust in the age of automated driving. In: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct. 2017. S. 15-21.
[18] Lulu Chang and Luke Dormehl. Our list of self-driving accidents, including the 'avoidable' Arizona Uber crash https://www.digitaltrends.com/cool-tech/most-significant-self-driving-car-crashes/ June 22, 2018 last accessed: March 2021
[19] Wagner, Alan R.; Borenstein, Jason; Howard, Ayanna. Overtrust in the robotic age. Communications of the ACM, 2018, 61. Jg., Nr. 9, S. 22-24.
[20] Parasumaran, Raja; Riley, Victor. Humans and automation: Use, misuse, disuse, abuse. Human factors, 1997, 39. Jg., Nr. 2, S. 230-253.
[21] Norman, Donald A. The 'problem'with automation: inappropriate feedback and interaction, not 'over-automation'. Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 1990, 327. Jg., Nr. 1241, S. 585-593.
[22] Ekman, Fredrick; Johansson, Mikael; Sochor, Jana. Creating appropriate trust in automated vehicle systems: A framework for HMI design. IEEE Transactions on Human-Machine Systems, 2017, 48. Jg., Nr. 1, S. 95-101.
[23] Endsley, M. Toward a theory of situation awareness in dyanmic systems. Human Factors, 1995, 37. Jg., S. 32-64.
[24] Dey, Debargha, et al. Distance-dependent ehmis for the interaction betIen automated vehicles and pedestrians. In: 12th international conference on automotive user interfaces and interactive vehicular applications. 2020. S. 192-204.
[25] Holländer, Kai. Applying the user-centered design process to external car displays. Mensch und Computer 2018-Workshopband, 2018.
[26] Lee, John D.; SEE, Katrina A. Trust in automation: Designing for appropriate reliance. Human factors, 2004, 46. Jg., Nr. 1, S. 50-80.
[27] Li, Jingyi; Holländer, Kai; Butz, Andreas. Introducing automated driving to the generation 50+. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings. 2019. S. 375-380.
[28] Ranasinghe, Champika, et al. Autonomous Vehicle-Pedestrian Interaction Across Cultures: Towards Designing Better External Human Machine Interfaces (eHMIs). In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020. S. 1-8.
[29] Endsley, Mica R., et al. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In: Proceedings of the human factors and ergonomics society annual meeting. Sage CA: Los Angeles, CA: SAGE Publications, 1998. S. 82-86.
[30] Holländer, Kai. Applying the user-centered design process to external car displays. Mensch und Computer 2018-Workshopband, 2018.
[31] Frison, Anna-Katharina, et al. Who is generation A? Investigating the experience of automated driving for different age groups. In: Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2018. S. 94-104.
[32] Gupta, Kunal, et al. In ai I trust: Investigating the relationship betIen biosignals, trust and cognitive load in vr. In: 25th ACM Symposium on Virtual Reality Software and Technology. 2019. S. 1-10.
[33] Inagaki, Toshiyuki; Itoh, Makoto. Human's overtrust in and overreliance on Advanced Driver Assistance Systems: a theoretical framework. International journal of vehicular technology, 2013, 2013. Jg.
[34] Payne, Stephen J. Users' mental models: The very ideas. HCI models, theories, and frameworks: Toward a multidisciplinary science, 2003, S. 135-156.
[35] Hart, Sandra G.; Staveland, LoIll E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Advances in psychology. North-Holland, 1988. S. 139-183.
[36] Liu, H., & Hiraoka, T. (2019, September). Driving behavior model considering driver's over-trust in driving automation system. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings (pp. 115-119).